Int. J. Production Economics 165 (2015) 260–272

Contents lists available at ScienceDirect

Int. J. Production Economics

journal homepage: www.elsevier.com/locate/ijpe

A big data approach for logistics trajectory discovery from RFID-enabled production data

Ray Y. Zhong [a,b,*], George Q. Huang [a], Shulin Lan [a], Q.Y. Dai [c], Chen Xu [d], T. Zhang [e]

[a] HKU-ZIRI Lab for Physical Internet, Department of Industrial and Manufacturing Systems Engineering, The University of Hong Kong, Hong Kong, China
[b] College of Information Engineering, Shenzhen University, China
[c] Guangdong Polytechnic Normal University, Guangzhou, China
[d] Institute of Intelligent Computing Science, Shenzhen University, Shenzhen, China
[e] Huaiji Dengyun Auto-parts (Holding) Co., Ltd., Huaiji, Zhaoqing, Guangdong, China

ARTICLE INFO

Article history:
Received 18 November 2013
Accepted 17 February 2015
Available online 23 February 2015

Keywords:
RFID
Big data
Logistics control
Trajectory pattern
Shopfloor manufacturing

ABSTRACT

Radio frequency identification (RFID) has been widely used in supporting the logistics management on manufacturing shopfloors where production resources attached with RFID facilities are converted into smart manufacturing objects (SMOs) which are able to sense, interact, and reason to create a ubiquitous environment. Within such environment, enormous data could be collected and used for supporting further decision-makings such as logistics planning and scheduling. This paper proposes a holistic Big Data approach to excavate frequent trajectory from massive RFID-enabled shopfloor logistics data with several innovations highlighted. Firstly, RFID-Cuboids are creatively introduced to establish a data warehouse so that the RFID-enabled logistics data could be highly integrated in terms of tuples, logic, and operations. Secondly, a Map Table is used for linking various cuboids so that information granularity could be enhanced and dataset volume could be reduced. Thirdly, spatio-temporal sequential logistics trajectory is defined and excavated so that the logistics operators and machines could be evaluated quantitatively. Finally, key findings from the experimental results and insights from the observations are summarized as managerial implications, which are able to guide end-users to carry out associated decisions.

1. Introduction

Big Data refers to a data set which collects large and complex data that is hard to process using traditional applications (Jacobs, 2009). With the increasing usage of electronic devices, our daily life is facing Big Data. For instance, taking a flight journey with A380, each engine generates 10 TB data every 30 min; more than 12 TB Twitter data are created daily and Facebook generates over 25 TB log data every day. It was reported that the per-capita capacity to store such data has approximately doubled every 40 months since 1980s (Manyika et al., 2011). Manufacturing and service industry largely involve in a range of human activities from high-tech products such as space craft to daily necessities like toothbrush. Manufacturing is regarded as the "hard" parts of economy using labors, machines, tools, and raw materials to produce finished goods for different purposes; while service sector is the "soft" part that includes activities where people supply their knowledge and time to improve productivity, performance, potential, and sustainability (Eichengreen and Gupta, 2013; Hill and Hill, 2009; Terziovski, 2010).

This paper is motivated by a real-life automotive part manufacturer which has used RFID technology for facilitating its shopfloor management over 10 years. Logistics within manufacturing sites like warehouse and shopfloors are rationalized by RFID so that materials' movements could be real-time visualized and tracked (Dai et al., 2012). The primary application of RFID for item visibility and traceability is rudimentary. First of all, estimation of delivery time on manufacturing shopfloor is basic for the sales department when getting a customer order. That helps to ensure the delivery date, which has been estimated from past experiences and time studies. Such estimation is not reasonable and practical given the difference of individual operators and seasonal fluctuation (e.g. peak and off seasons). Secondly, RFID-enabled real-time manufacturing, planning and scheduling on shopfloors heavily relie on the arrival of materials, thus, the decisions on logistics trajectory are critical. This company carries the decision using paper sheets manually which always make the material delay. That causes many replanning and rescheduling, which greatly affect the production efficiency. Finally, the space on the manufacturing shopfloor is limited. As a result, the logistics trajectories of materials should be optimized. Currently, the logistics is not

* Correspondence to: 8-23 Haking Wong Building, Pokfulam Road, Hong Kong, Tel.: +852 22194298; fax: +852 28586535.
E-mail address: zhongzry@gmail.com (R.-n. Zhong).

http://dx.doi.org/10.1016/j.ijpe.2015.02.014
0925-5273/

well-organized, which causes high WIP (Work-In-Progress) inventory on manufacturing shopfloors.

In order to address the above hurdles, the senior management made a decision to explore a solution from making full use of such RFID-enabled logistics Big Data. Unfortunately, they are facing several challenges. Firstly, manufacturing resources equipped with RFID devices are converted into smart manufacturing objects (SMOs) whose movements generate large number of logistics data since SMOs are able to sense, interact, and reason each other to carry out logistics logics. The enormous RFID-enabled logistics data closely relate to the complex operations on manufacturing shopfloors (Zhong et al., 2013). That leads to a great challenge for further analysis and knowledge discovery. Secondly, the RFID-enabled logistics Big Data usually include some “noise” such as incomplete, redundant, and inaccurate records, which could greatly affect the quality and reliability of decisions. Therefore, elimination of the redundancy is necessary (Zhong et al., 2013). However, current methods are not suitable for removing the above noises due to the high complex and specific characteristics of RFID Big Data. Finally, mining frequent trajectory knowledge is significant for determining the logistics plans and layout of distribution facilities. However, the knowledge hidden in the RFID-enabled Big Data is sporadic. That means hundreds of RFID records may create a piece of information which indicates the detailed logic operations. To achieve the creation is very challenging.

This paper proposes a holistic Big Data approach to excavate the frequent trajectory from massive RFID-enabled manufacturing data for supporting production logistics decision-makings. This approach comprises several key steps: warehousing for raw RFID data, cleansing mechanism for RFID Big Data, mining frequent patterns, as well as pattern interpretation and visualization.

The rest of this paper is organized as follows. Section 2 briefly reviews the related work such as RFID in production logistics control, frequent trajectory pattern mining, and Big Data in Manufacturing. Section 3 presents a RFID-enabled logistics control through introducing the deployment of RFID devices to create a RFID-enabled ubiquitous manufacturing site and logistics operations within it. Section 4 demonstrates the RFID logistics data warehouse and spatio-temporal sequential RFID patterns. Section 5 proposes a Big Data approach in terms of framework, key algorithms for discovering trajectory knowledge from RFID-enabled manufacturing data, as well as an example to validate the proposed approach. Experiments and discussions, including design of experiments, evaluations, and managerial implications are presented in Section 6. Section 7 concludes this paper by giving our major findings and future work.

2. Literature review

This section reviews related research which is categorized into three dimensions: RFID in production logistics control, frequent trajectory pattern mining, and Big Data in manufacturing.

2.1. RFID in production logistics control

Due to the bright advantages of RFID technology, it has been widely used for production and logistics control in supply chain management (SCM) (Sarac et al., 2010). This section briefly reviews this topic from theoretical and practical aspects.

In theoretical perspective, large number of models and frameworks has been proposed. For creating value from RFID-enabled SCM, a contingency model was proposed in logistics and manufacturing environments (Wamba and Chatfield, 2009). The model draws on a framework and analyzes five contingency factors which greatly influence value creation. Since RFID could be used for supporting different decision-makings, theoretical models are important. A cost of ownership (COO) model for RFID logistics system was introduced in order to support the decision-making process in an infrastructure construction (Kim and Sohn, 2009). This paper established three scenarios using the RFID system to evaluate the expected profit, helping companies to choose the most beneficial RFID logistics system. RFID is supposed to facilitate end-users decision-making in production logistics control. To assist the managers' determination of appropriate operational and environmental conditions under the adoption of RFID, a framework was presented at different levels of collaboration through a comprehensive simulation model (Sari, 2010). Within the RFID-enabled environment, real-time data could be captured and collected. These data can be used for different purposes. A model thus for determining the RFID real-time information sharing and inventory monitoring works on environmental and economic benefits was proposed (Nativi and Lee, 2012). This study implies that the economic benefits are achieved through carrying out numerical studies. In practical perspectives, RFID technology has been used for controlling the production and logistics. A warehouse management system (WMS) with RFID was designed for monitoring resources and controlling operations (Poon et al., 2009). In this system, the data collection and information sharing are facilitated by RFID. With the information, case-based logistics control is realized. In order to improve remanufacturing efficiency, RFID technology was used for examining the benefits in practice (Ferrer et al., 2011). This paper gives a framework for considering the RFID adoption in terms of location identification and remanufacturing process optimization. Currently, autonomy in production and logistics attracts many attentions in practical fields. RFID was investigated to autonomous cooperating logistics processes to react quickly and flexibly to an increasing dynamic ambience (Windt et al., 2008). This paper evaluates the feasibility and practicality by means of an exemplary shopfloor scenario. The fast-moving consumer goods (FMCG) supply chain with RFID was quantitatively assessed within a three-echelon SCM, which contains manufacturers, distributors, and retailers (Bottani and Rizzi, 2008). RFID technology adoption with pallet-level tagging, from this research, shows that positive revenues for all supply chain stakeholders could be achieved; while, a case-level tagging will add costs for manufacturers, resulting in negative economical results.

Cases with RFID application in production and logistics control from practical aspects are also widely studied and reported. Eastern Logistics Limited (ELL), a medium-sized 3 PL company used RFID technology in visualizing logistics operations (Chow et al., 2007). This case shows the enhanced performance of its supply chain partners in reduced inventory level, improved delivery efficiency, and avoidance of out-of-stock. In order to study the factors influencing the use of RFID in China, 574 logistics companies were analyzed in terms of technological, organizational, and environmental aspects (Lin and Ho, 2009). Most of the cases reveal the advantages of using RFID for dealing with data capturing in the initial stage. After the data collection, further applicable dimension is explored like visibility and traceability. A manufacturing services provider company was introduced for assessing the RFID deployment at one of its production line for tracing components (Chongwatpol and Sharda, 2013). After the RFID deployment, the cycle time, machine utilizations, and penalty costs are significantly improved by comparing the RFID-based scheduling and traditional approach. For examining the impact of RFID-enabled supply chain on pull-based inventory replenishment, a case study in TFT-LCD (Thin-film-transistor liquid-crystal display) industry was illustrated (Wang et al., 2008). From this case, it is observed that the total inventory cost could be cut down by 6.19% by using the RFID-enabled pull-based supply chain. More real-life cases using RFID for supporting real-time production, logistics control and supply chain management could be found from (Dai et al., 2012; Ngai et al., 2008; Sarac et al., 2010; Zhong et al., 2014).